REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-10-0334

1. REPORT DATE (DD-MM-YYYY) Oct 25, 2006	2. REPORT TYPE Final Technical Report	3. DATES COVERED (From - To) Dec 15, 2002 - Jun 14, 2006
4. TITLE AND SUBTITLE Higher Resolution Spectral Analysis		5a. CONTRACT NUMBER 5b. GRANT NUMBER F49620-03-1-0120 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Tryphon T. Georgiou		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Minnesota Dept. of Elect & Comp. Eng. 200 Oak Street Minneapolis, MN 55455-5200		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research (AFOSR) Attn: Lt. Col. Scott Wells 875 North Randolph Street, Arlington, VA 22203		10. SPONSOR/MONITOR'S ACRONYM(S) AFOSR 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

N/A

14. ABSTRACT

We have developed theory and software for multivariable subspace analysis and for identifying emission spectral lines in multivariable time-series. This is expected to have an impact in polarimetric radar, where the underlying energy distribution is truly multivariable. Finally we have introduced intrinsic notions of distance between power spectra. These measures quantify the degradation of performance when a wrong choice is made between two alternative.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Tryphon T. Georgiou
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (include area code) (612) 625-3303

Final Report to the
Air Force Office for Scientific Research
Grant Number: # AF/F49620-03-1-0120
Title: High resolution spectral analysis

Principal investigator:
Tryphon T. Georgiou
University of Minnesota

October 27, 2006

Final Report to the
Air Force Office for Scientific Research
Grant Number: # AF/F49620-03-1-0120

1. **Date submitted:** October 27, 2006

2. **Title:** High Resolution Spectral Analysis.

3. **Legal institution name:**
Unversity of Minnesota
Office of Sponsored Projects Administration
200 Oak street, S.E., suite 450
Minneapolis, MN 55455

4. **Principal investigator:**
Tryphon Georgiou
Electrical & Computer Engineering
(612) 625-3303, tryphon@umn.edu

5. **Business official:**
Office of Sponsored Projects Administration
200 Oak street, S.E., suite 450
Minneapolis, MN 55455
(612) 624-5599, awards@tc.umn.edu

Chapter 1

Abstract

The research conducted with AFOSR support under Grant # AF/F49620-03-1-0120 aimed at developing algorithms and theory for reliable retrieval of information hidden in noisy measurements.

A wide range of technologies, from medical diagnostics to reconnaissance and targeting, rely critically on the quality of available data. Our ability to control is often only hindered by our ability to see. The data, whether collected using ultrasound transducers, radar, or a distributed array of sensors have one thing in common, the useful information is often swamped in noise. The research focused on developing a next generation of spectral analysis tools and resolution standards that provide the maximal amount of useful information as well as quantitative assessment of the remaining uncertainty.

The research effort has generated algorithms that have a built-in ability to focus in on particular futures of recorded signals. The project has undergone several phases already. Early work, in collaboration with Professors C.I. Byrnes and A. Lindquist, produced a method and apparatus for a Tunable High-REsolution spectral Estimator (U.S. patent No. 6,400,310) nicknamed THREE. A very substantial improvement in resolution over prior state-of-the art was documented via theoretical as well as experimental studies. A number of specialized algorithms spawned from this and have since been tested on synthetic aperture radar imaging using MSTAR data and on ultrasound imaging in collaboration with Professor E. Ebbini of the University of Minnesota.

Recently, in collaboration with Dr. Dan Herrick (AFRL/DESA), high resolution methods are being tailored to, disturbance isolation of a targeting system (e.g., laser) using input from a distributed array of

sensors. High resolution methods can be used for sifting efficiently through huge amounts of data so as to identify sources and directionality of disturbances. Thereby it is anticipated that the research completed under the grant will further contribute to critical targeting technologies.

Besides the development of algorithms, the research team has invested on the development of mathematical tools and standards for quantifying resolution and reliability in spectral analysis and in estimation. These are expected to guide the tuning of control and measurement strategies in the near future.

Contents

Chapter 2

Summary of Research Contributions

The research conducted under the present grant (# AF/F49620-03-1-0120) has led to a range of tools for robust and high resolution signal analysis. Broadly, these tools consist of

(i) new theoretical results for signal analysis,

(ii) numerical algorithms for high resolution analysis, and

(iii) quantitative metrics for assessing performance.

Moreover, under the present grant, application studies have been conducted in utilizing these techniques in the context of sensor networks, ultrasound imaging, and synthetic aperture radar imaging.

Signal processing is enabling technology for a broad range of applications, from medical diagnostics to reconnaissance and communications. Advances in signal analysis techniques often have an immediate impact on the state-of-the art in such application areas. The focus of the research has been an apparent need for robust, flexible, and high resolution analysis tools. The hallmark of our approach has been the development of algorithms with a built-in ability to focus in on particular futures of recorded signals.

The approach we initiated has undergone several phases already. Early work, in collaboration with Professors C.I. Byrnes and A. Lindquist, produced a method and apparatus for a Tunable High-REsolution spectral Estimator (U.S. patent No. 6,400,310) nicknamed THREE, which has been especially suitable for time-series analysis and speech analysis. A number of specialized algorithms spawned from this. These algorithms allow integrating data from a variety of sensors with arbitrary geometry and originating from multivariable time-series. The ability to deal with arbitrary sensor geometries and, at the same time, to ensure theoretical performance bounds, represents a quantum jump in the state-of-the art.

The signal analysis tools that we have developeded are intended to separate frequency components, identify deterministic components from random ones, and to detect long range vs. short range interactions. The components obtained in such an analysis stage, help identify a wide range of underlying physical causes, and dynamical dependencies. A particular advantage of these new methods is that they have a built-in mechanism for taking into account prior information about the processes under consideration. On the application side, we have explored the relevance of these techniques in two main areas. We conducted case studies focusing on measuring the temperature of (artificial) tissue in a non-invasive manner by analysing ultrasound echo. The purpose of such "non-invasive sensing" is to provide reliable measurement of tissue temperature for computer guided tumor ablation and therapy [31]. The results have been very encouraging and substantially better than earlier state-of-the art. We have also explored the use of such high resolution techniques in synthetic apperture radar (SAR) with similar results.

A most significant breakthrough has been a theory for analyzing and integrating data from distributed sensor arrays (e.g., see [9]). Interestingly, this advancement shares the same basic framework with our techniques for high resolution spectral analysis. In either, we seek a power distribution which is consistent with measurements. Pre-conditioning of the data may be taken into account, and post-processing can be tailored to the task at hand based on prior information (e.g., accounting for known portion of the power and tuning for high resolution in a particular frequency or spatial sector). Identifying power distributions consistently with given measurements is treated as an inverse problem. The family of such distributions is suitably parametrized, and the size of the family represents a measure of uncertainty. When data are collected via a spatially scattered collection of sensors, the relevant imaging and sensor analysis problems are cast as moment problems.

Typically, data and measurements represent radar/sonar echo, a speech recording, etc., and may be sampled non-uniformly with gaps on the record. Our computational theory allows solving the most general such multivariable and multidimensional moment problems by providing representative spectra, as well as a complete description of all spectra which are consistent with the data. Besides the relevance of these techniques in analysis, they also impact on the design and optimal distribution of sensors.

The publications listed below document our research and accomplishments. Our joint work with C. Byrnes and A. Lindquist [3], where foundations of our framework were first laid out, received the G.S. Axelby outstanding paper award from the IEEE Control Systems Society in 2003, and a U.S. patent [41] which was based on this and subsequent work. We mention that earlier joint work of the PI with Professor M.C. Smith (University of Cambridge) on metrics for robust control analysis and synthesis, work which was also supported by AFOSR, received the G.S. Axelby outstanding paper award twice. First in 1992, for a robust control theory and design tools for linear systems, and then again in 1999 for robust control theory and tools for nonlinear systems. The recent work under the current grant, on robust high resolution spectral analysis, can be broadly classified into the following categories with some overlaps.

2.1 High resolution analysis and applications

Publications: [31, 36, 32, 11, 22, 6, 23, 8, 13, 14, 15, 34, 30, 17, 18, 40, 37]

Our framework for spectral analysis was initiated in [19, 20, 21]. It was influenced by earlier joint collaborative work of the PI with Chris Byrnes and Anders Lindquist in [3, 4]. This also led to a U.S. patent [41] on tunable high-resolution spectral estimators. Publication [32] explores basic tradeoffs between resolution and robustness of such estimators, and outlines how to tune these for optimal performance. In [36, 37, 40, 35] we explain advantages of the new techniques and insight in antenna arrays, SAR, and in multi-rate signal processing. In [31, 34] we demonstrated the use of our new methods in non-invasive ultrasound temperature sensing. This work is on-going. The goal is to use non-invasive sensing for computer controlled tumor ablation. In [13, 14, 15] we pointed out that a key step in spectral analysis, the step of estimating statistics, may not provide data which are consistent with underlying dynamics. In this case, resolution can be dramatically improved if care is taken to adjust the statistics so as to conform with known underlying dynamics (in a way which extends the celebrated contribution by Burg [27] to the case of multivariable processes). Publications [30, 23, 11, 22, 18] deal with assessing the level of spectral uncertainty, and then presenting canonical decompositions for use in spectral analysis problems.

2.2 Multi-variable & multi-dimensional moments

Publications: [9, 11, 22, 8, 12, 13, 14, 17]

In these publications we solve the most general multi-variable and multi-dimensional moment problem. In a very general sense, the data for modeling, identification, spectral analysis, etc. amount to moment constraints on a power density function which is possibly multivariable (matrix-valued) and multidimensional (spatio-temporal). Our theory in [9] gives a way to determine and parametrize all consistent distributions. Publications [22, 11] develop further a very important "boundary" case of singular data sets. Publications [12, 13, 14] are mostly on a static but multivariable version of such problems and corresponding numerical issues.

The importance of this stems from the fact that it encompasses the most general situation where data is available from a distributed array of sensors. The sensors do not need to collect the data in synchronized manner. Also the data may originate in a spacio-temporal and multi-variable distribution of scattered power. Our framework allows for a computational theory for integrating data collected by such distributed arrays of sensors for the purpose of identifying the distribution of power and properties of the underlying scattering field. Such problems are encountered in a wide range of applications which includes radar technology, ultrasound imaging, and others.

2.3 Analytic interpolation with degree constraint

Publications: [5, 22, 8, 9, 17]

The problem of analytic interpolation with degree constraint was introduced in the PI's work in the early 1980's (1983 Ph.D. thesis). Important contributions by Chris Byrnes and Anders Lindquist re-kindled interest in the problem, and in the joint work (together with A. Megretski) [5] a rather complete theory for the (scalar version) of the problem was finally completed. Interest in this problem stems from applications in control and signal processing. The theory in [22, 8, 9, 17] is relevant in addressing the multivariable version of the problem (which is essential for multivariable control applications). Work on the multivariable problem is still in progress along the lines of [9] and will appear shortly. This work entails a complete parametrization of all solutions to a multivariable Nehari type of interpolation problem, which have a Macmillan degree less than or equal to the generic degree prescribed by the problem data.

2.4 Feedback control design

Publications: [44, 10, 16, 39]

Reference [44] deals with the numerical solution of certain optimal periodic control problems. These arose in studies in periodic drug delivery and in chemical engineering applications. References [10, 16, 39] deal with applying our theory on interpolation with degree constraint to controller design with dimensionality constraints. Work in [10] compares our framework for controller design with degree constraints to an alternative which is based on solving linear matrix inequalities.

2.5 Metrics for spectral uncertainty

Publications: [24, 25]

Despite the centrality of spectral analysis in a wide range of scientific disciplines, no agreement exists as to what an appropriate distance measure between spectral density functions is. Some of the key contenders have been the Bregman distances, the Kullback-Leibler-von Neumann distance, the Itakura-Saito distance, and finally Battacharrya and Mahalanobis-type variants. Certain of these distances have a definite relevance when used to discriminate between two probability density functions. Yet ***none*** has any meaningful interpretation when applied to power spectral density functions.

In fact, surprisingly, the question on how to quantify uncertainty in spectral analysis has received little attention in the past. The development of suitable metrics for quantifying distances between power spectra has been a focus in our research. These metrics are essentials tool for assessing uncertainty, robustness, and performance of signal analysis techniques. Our formalism mimics ideas from Information geometry which was invented so as to quantify uncertainty in statistical inference. Information geometry endows probability distributions with a natural (Fisher infromation) metric. We similarly endow density functions with a metric with a natural interpretation based on prediction theory. We developed new distance measure between power spectral densities [24, 25]. These metrics are the first with any clear interpretation. The relevant notion of distance quantifies differences in predictability properties between respective random processes. It can be used to detect subtle changes in the spectral content of nonstationary signals and, thereby, effectively identify drifts, transitions, and events.

Chapter 3

Technical Highlights

The subject of our research relates to modeling and identification of multivariable and multidimensional time-series with a focus on sensor arrays. The focus has been an apparent need for robust, flexible, and high resolution analysis tools. The hallmark of our approach has been the development of algorithms with a built-in ability to focus in on particular futures of recorded signals. The intended applications include identification, array signal processing, and data mining.

In Section 3.1, we overview methods for multivariable spectral analysis based on generalized statistics. These inherit superior resolution properties from utilizing generalized statistics and represent a generalization of our earlier Tunable High REsolution Estimator (THREE) reported in [4]. We discuss additive decompositions of covariance statistics and their significance in identification, and outline basic theory which underlies our advances on high resolution identification and spectral analysis [3, 4, 6, 7, 8, 9, 19, 21, 22, 31, 32].

In the backdrop of robust control and of the gap metric (introduced by Zames and El-Sakkary and developed during the past decade by the PI in collaboration with M.C. Smith), the PI seeks quantitative measures of signal affinity as well as objective distances between signal statistics. Traditional vector metrics (e.g., Euclidean) are not satisfactory as distances between statistical quantities since they do not acknowledge their structure as a positive cone. Alternative entropy-based metrics are sought and a natural intrinsic metric between spectral density functions is discussed in Section 3.2. The purpose is to quantify uncertainty and resolution in signal analysis, to assess robustness of signal processing algorithms, and to facilitate correlation analysis in large databases.

Part of the motivation and impetus for the research has been provided by our recent collaborative work on two fronts: Non-invasive sensing for medical applications, and distributed sensing and control for vibration isolation. Measuring heart conditions such as mitral regurgitation or measuring temperature inside tissue for purposes of computer guided tumor ablation or therapy, in a non-invasive manner, requires exceedingly high resolution and robustness (see e.g., [31]). Another ubiquitous situation arises when a large collection of signals becomes available (e.g., via a distributed sensor array or in a large database) and fast and high resolution correlation analysis is sought to identify affinity, coherence, and relevance of various signals. Such is the task when we seek to identify sources of excitations in distributed media, the origin and pathway of disturbances in spatial structures, for the purposes of modeling/prediction/control of distributed physical systems —a model application being the disturbance isolation of a targeting system (e.g., a laser) using feedback from measurements collected using a distributed sensor array. Similar tasks are pertinent in signal classification and system identification, in general.

On the technical side, the research reported herein has led to reliable methods for high resolution spectral analysis of multivariable processes, as well as to distance measures for quantitative assessment of uncertainty and of resolution in signal analysis applications. In the last section 3.3, as an example, we outline certain Matlab-based routines and their use for high resolution analysis of scalar time-series. There routines are available the PI's website.

3.1 Generalized statistics & resolution

Our first step is to explain what is meant by "generalized statistics" and why is this concept important. Consider a stationary random process $\{u_k\}$ with zero mean and power spectral density $f_u(\theta)$. The autocorrelation samples $R_k = \mathcal{E}\{u_\ell u_{\ell+k}\}$ of u_k are the Fourier coefficients of $f_u(\theta)$, i.e.,

$$R_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-j\theta} f_u(\theta) d\theta, \text{ for } k = 0, \pm 1, \pm 2, \ldots,$$

$R_{-k} = R_k^*$, while $f_u(\theta)$ is given by the Fourier series $\sum_{k=-\infty}^{\infty} R_k e^{-jk\theta}$. Occasionally, $\{u_k\}$ is not directly observable in which case one may not be able to estimate autocorrelation samples. For instance, if $x_k = ax_{k-1} + u_k - bu_{k-1}$ is a first-order system $(-1 < a < 1)$ and if only $\{x_k\}$ is available, then it is natural to estimate statistics of $\{x_k\}$ instead. These statistics represent moments of $f_u(\theta)$ with respect to kernel functions which differ from $e^{jk\theta}$. For example, the variance of x_k,

$$\mathcal{E}\{x_k^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\frac{e^{j\theta} - b}{e^{j\theta} - a}|^2 f_u(\theta) d\theta$$

is a moment of $f_u(\theta)$ with respect to the kernel function $|(e^{j\theta} - b)/(e^{j\theta} - a)|^2$. Such filtering may inherently be part of a measuring apparatus, but it may also be introduced to improve S/R and resolution as discussed in e.g, [32, 20].

Thus, in general, it is customary to refer to any moments of $f_u(\theta)$ as generalized statistics of the underlying random process. Not all such moments originate in ordinary time-filtering, and not all correspond to rational kernel functions. In fact, a most challenging and very common situation arises when the indexing in $\{u_k\}$ refers to space and not time. Take for instance an array of sensors with three elements, linearly spaced at distances 1 and $\sqrt{2}$ wavelengths from one another, and assume that (monochromatic) planar waves, originating from afar, impinge upon the array. This is exemplified in Figure 3.1. Assuming

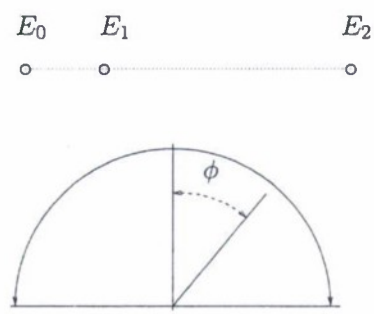

Figure 3.1: Non-equispaced sensor array

that the sensors are sensitive to disturbances originating over one side of the array, with sensitivity independent of direction, the signal at the

ℓth sensor is typically represented as a superposition

$$u_\ell(t) = \int_0^\pi A(\theta) e^{j(\omega t - p x_\ell \cos(\theta) + \phi(\theta))} d\theta,$$

of waves arising from all spatial directions $\theta \in [0, \pi]$, where ω is the angular time-frequency (as opposed to "spatial"), x_ℓ the distance between the ℓth and the 0th sensor, p the wavenumber, and $A(\theta)$ the amplitude and $\phi(\theta)$ a random phase of the θ-component. Typically, the phase $\phi(\theta)$ for various values of θ are uncorrelated. The term $p x_\ell \cos(\theta)$ in the exponent accounts for the phase difference between reception at different sensors. For simplicity we assume that $p = 1$ in appropriate units. Correlating the sensor outputs we obtain

$$R_k = E\{u_{\ell_1} \bar{u}_{\ell_2}\} := \int_0^\pi e^{-jk\cos(\theta)} f(\theta) d\theta \tag{3.1}$$

where $f(\theta) = |A(\theta)|^2$ now represents *power density*, and $k = \ell_1 - \ell_2$ with $\ell_1 \geq \ell_2$ and belonging to $\{0, 1, \sqrt{2} + 1\}$ (k is kept as a "non-integer" index in R_k from mnemonic purposes). Thus,

$$k \in \mathcal{I} := \{0, 1, \sqrt{2}, \sqrt{2} + 1\}. \tag{3.2}$$

The only significance of our selection of distances between sensors, that gave rise to the rather unusual indexing set (3.2), is to underscore that there is no algebraic dependence between the kernel functions

$$1,\ e^{-j\cos(\theta)},\ e^{-j\sqrt{2}\cos(\theta)},\ e^{-j(\sqrt{2}+1)\cos(\theta)}.$$

Even more challenging situations arise when (i) the kernel functions represent Green's functions or transfer functions in a general spatial domain, e.g., in case sensors are scattered in a random pattern in $\mathbb{R}^3$, and (ii) when statistics are obtained from observations which are non-equispaced in time (also, *random sampling*).

Let us revisit the situation of ordinary autocorrelation samples. Given a finite observation record $\{u_0, \ldots, u_N\}$ we typically estimate a finite sequence $\{R_0, R_1, \ldots, R_n\}$ (e.g., via sample averaging). The Toeplitz matrix

$$T_n := \begin{bmatrix} R_0 & R_{-1} & \ldots & R_{-n} \\ R_1 & R_0 & \ldots & R_{-(n-1)} \\ \vdots & & & \vdots \\ R_n & R_{n-1} & \ldots & R_0 \end{bmatrix}$$

is nonnegative definite since $f_u(\theta) \geq 0$ and since

$$T_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} G(e^{j\theta})f_u(\theta)G(e^{j\theta})^* d\theta,$$

where

$$G(e^{j\theta}) = \left[\begin{array}{cccc} 1 & e^{j\theta} & \dots & e^{jn\theta} \end{array}\right]',$$

and "$'$" denotes transpose while "$*$" denotes complex conjugate transpose. The (column) "Fourier vector" $G(e^{j\theta})$ is referred to as a Fourier vector, and as θ varies, it defines a curve in a complex space which is known as the "array manifold". Non-negativity of T_n turns out to be sufficient for the existence of a power spectral density which is consistent with the moments in T_n. There is a rather rich theory on how much T_n is telling us about the power spectum, and how to reconstruct representative spectra (maximum-entropy, etc.) which are consistent with the partial sequence of autocorrelation statistics. This goes back to the theory of the trigonometric moment problem and of orthogonal polynomials, and forms the basis of the so-called "modern nonlinear spectral analysis methods" [27]. An alternative way to reconstruct $f_u(\theta)$, based on T_n, is the periodogram/correlogram

$$\hat{f}(\theta) \ := \ \frac{1}{n+1}G(e^{j\theta})^* T_n G(e^{j\theta}). \tag{3.3}$$

This is an approximation of $f_u(\theta)$ — see [38], as is the Capon (or maximum-likelihood) estimate of $f_u(\theta)$ given by

$$\hat{f}(\theta) := \frac{1}{n+1}\left(G(e^{j\theta})^*(T_n)^{-1}G(e^{j\theta})\right)^{-1}. \tag{3.4}$$

Weighted versions of the autocorrelation coefficients can be used in order to trade-off resolution with robustness. I.e., using various windowing functions w_k (Hamming, Kaiser, etc.) one may replace R_k with $R_k w_k$ in the above. These ideas are classical, were extensively studied decades ago, and remain the workhorse of signal analysis applications to this day. Yet, it is a striking fact that a multivariable version of such successful tools has largely been absent (i.e., a periodogram-like method for inherently multivariable processes). In our software tools we generalize and take advantage of similar ideas in the context of generalized statistics, replacing the state covariance R by Hadamard products with suitable windowing matrices.

A further striking fact is that the corresponding issues when $G(e^{j\theta})$ is not an ordinary Fourier vector, have not been studied with the exception of the somewhat ad-hoc beamspace techniques. The recent work by the PI (see [9, 8, 40]) has addressed such issues *on a firm theoretical basis.* For instance, returning to the example of the non-equispace antenna array in Figure 3.1 and R_k for $k \in \mathcal{I}$ in (3.1), it is important to determine whether estimated values for the moments are consistent with the geometry of the array, and if so to characterize all consistent power spectra. In the present situation (packaging R_k's in (3.1) into a matrix and setting $\tau = \cos(\theta)$) the nonnegativity of

$$R := \int_{-1}^{1} \begin{bmatrix} 1 \\ e^{-j\tau} \\ e^{-j\sqrt{2}\tau} \end{bmatrix} \frac{f(\cos^{-1}(\tau))}{\sqrt{1-\tau^2}} \begin{bmatrix} 1 & e^{j\tau} & e^{j\sqrt{2}\tau} \end{bmatrix} d\tau$$

which, in the obvious indexing turns out to be

$$R = \begin{bmatrix} R_0 & R_1 & R_{\sqrt{2}+1} \\ \bar{R}_1 & R_0 & R_{\sqrt{2}} \\ \bar{R}_{\sqrt{2}+1} & \bar{R}_{\sqrt{2}} & R_0 \end{bmatrix}, \tag{3.5}$$

is only a necessary condition. The fact that it is *not* sufficient (see e.g., [7, page 786]) motivated our recent work and *led to a rather complete answer/theory* documented in [9, 8].

We now highlight some of the important findings. First we deal with the case where the "Fourier vector" $G(e^{j\theta})$ is replaced by the transfer function of a linear, time-invariant, discrete-time (input-to-state) dynamical system

$$x_k = Ax_{k-1} + Bu_k, \text{ with } k \in \mathbb{Z},$$

x_k being the state-vector and A, B matrices in $\mathbb{R}^{n\times n}$ and $\mathbb{R}^{n\times m}$, respectively. Then the input-to-state transfer function $G(e^{j\theta}) = (I - e^{j\theta}A)^{-1}B$ is matricial and the random process $\{u_k\}$ vectorial. If u_k is white noise (with covariance matrix $Q \geq 0$), then it is well known and easy to see that the state covariance

$$R := \mathcal{E}\{x_k x_k'\}$$

satisfies the Lyapunov equation $R - ARA' = BQB'$. However, surprisingly, the case where u_k is not white was dealt only recently (under

AFOSR support by the PI in [20, 21]). The correspondence between R, A, B and input power spectra f_u is detailed in [20, 21, 22]. Briefly, a state covariance for the above system satisfies

$$\operatorname{rank}\begin{bmatrix} R - ARA^* & B \\ B^* & 0 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} \tag{3.6}$$

where 0 is the zero matrix of appropriate dimension. An alternative characterization amounts to the solvability of

$$R - ARA' = BH' + HB'$$

for a matrix H which is of the same size as B. Conversely, provided R satisfies either of the above two equivalent conditions, and provided it is non-negative definite, there exists a power spectrum for a candidate input that gives rise to such state-statistics (this was shown in [20]). The parametrization of all consistent power spectra and related computational issues has been the subject of [20, 21, 22]. The relevant realization theory for matricial power spectral densities amounts to analytic interpolation with positive-real matricial functions and thus, echoes a lot of the usual tools and constructions in H_∞-control theory.

The motivation for considering state-covariances of linear systems, was to develop a theory for high resolution spectral analysis following [3, 4, 19]. Our joint work with C. Byrnes and A. Lindquist led to a U.S. patent [41]. The main idea in [3] arose from the simple observation that the autocorrelation samples of a time-series correspond to interpolation conditions for a positive-real function related to the power spectrum, at the origin. In some detail, if $R_k = \mathcal{E}\{u_\ell u_{\ell+k}\}$ as before, then the power spectral density $f_u(\theta)$ is simply the real part of the positive real function

$$F(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1+ze^{j\theta}}{1-ze^{j\theta}}f(\theta)d\theta = R_0 + 2R_1 z + 2R_2 z^2 \ldots.$$

Thus, the R_k's relate to the value of $F(z)$ and its derivatives at the origin. This fact generalizes to cases where the statistics are taken at the state or output of any dynamical system. For instance, if $x_k = ax_{k-1} + u_k$ is a first-order system and $-1 < a < 1$ then

$$\mathcal{E}\{x_k^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{|e^{j\theta}-a|^2}f(\theta)d\theta = \frac{1}{1-a^2}F(a)$$

from which we readily obtain an interpolation constraint on $F(z)$ at $z = a$. In general, intuitively, superior resolution is achieved by selecting data-dependent interpolation constraints at points proximal to the unit-disc sector of a targeted frequency band. In general, the filter may reflect dynamics of sensors but it can also be virtual, focusing on the frequency range of interest. Given interpolation constraints for the power spectrum, a whole range of tools of the nonlinear methods [27] extends to this framework (encompassing so-called beamspace techniques in antenna arrays). The design of input-to-state filters and relevant tradeoffs between robustness and resolution have been addressed in [32], and is part of continuing research and development of algorithms.

We now highlight the case where the "Fourier vector" is replaced by a Green's/transfer function $G(e^{j\theta})$ with no apparent shift structure. Turning once more to the non-equispaced antenna array in Figure 3.1, we seek a power density function $f(\theta)$ consistent with the statistics which is closest to a "prior" $f_{\text{prior}}(\theta)$ in the sense of, say, a Kullback-Leibler distance

$$\mathbb{S}(f||f_{\text{prior}}) := \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(f_{\text{prior}}\log(f_{\text{prior}}) - f_{\text{prior}}\log(f)\right) d\theta.$$

The minimizing solution can be written in closed form

$$f(\theta) = \frac{f_{\text{prior}}(\theta)}{\text{Re}\{\lambda_o G(e^{j\theta})\}}$$

where λ_o denotes a (row) vector of Lagrange multipliers for the minimization problem. These multipliers can be easily computed so that $f(\theta)$ abides by the given statistics, provided of course that the statistics are consistent with the structure of $G(e^{j\theta})$. A homotopy method was proposed in [8, 9] leading to a differential equation for $\lambda(\tau)$ in a homotopy variable τ. If the statistics are consistent with the structure of $G(e^{j\theta})$, then $\lambda(\tau) \to \lambda_o$ as $\tau \to 1$, otherwise $\lambda(\tau)$ escapes to ∞. The rôle of f_{prior} is to introduce prior information, but can also be used to parametrize all solutions to the moment problem, since choices of f_{prior} lead to the complete set of f's such that

$$R = \frac{1}{2\pi}\int_{-\pi}^{\pi} G(e^{j\theta})f(\theta)G(e^{j\theta})^{*}d\theta. \tag{3.7}$$

We would like to emphasize that the theory in [9] applies to the case of matricial power density functions f_u (e.g., spectral density functions of multi-variable processes), as well as to cases where the support is multi-dimensional (e.g., space-time distributions, or $\theta \in \mathbb{R}^\ell$ with $\ell > 1$ in general) in which case the integrals are interpreted accordingly.

The observation that singularities in a covariance matrix reveal deterministic linear dependences between observed quantities, forms the basis of a wide range of techniques, from Gauss' least squares, to principal component analysis (PCA, GPCA), to modern subspace methods in time-series analysis. This observation suggests that a decomposition of covariance data into "signal + noise," in accordance with a suitable postulate, leads to identification of such deterministic dependences.

We first discuss the implications of the observation in time-series analysis. Here, traditionally, one seeks a white-noise component of maximal variance which is consistent with estimated statistics. For instance, if T_n represents the $(n+1) \times (n+1)$ Toeplitz matrix formed out of the first $n+1$ autocorrelation samples of a scalar random process, the minimal eigenvalue $\lambda_{\min}(T_n)$ of T_n represents the maximal power of white noise which is consistent with this autocorrelation data. Furthermore, $T_n - \lambda_{\min}(T_n)I$ is singular and corresponds to a deterministic random process made up of at most n-complex sinusoidal components. This fact (albeit in a different language) was already known to Carathéodory and Fejér in the early part of the 20th century, and was used by them to show that positivity of T_n is sufficient for the solution of the relevant trigonometric moment problem. It was recognized by Pisarenko in the 1960's for its relevance in signal analysis and this fact forms the basis of certain widely used high resolution methods for spectral analysis known as MUSIC (MUltiple SIgnal Classification) and ESPRIT (EStimation of Parameters by Rotational Invariant Techniques) —see [7, 19, 38].

Despite being widely used, no multivariable generalization of the Carathéodory-Fejér-Pisarenko decomposition had been devised, until the PI's recent work [22] under the current AFOSR grant. In [22] we have shown that a direct multivariable analog is not possible. More specifically, we have shown that after we account for white noise of maximal power consistent with the data, the remaining variance cannot be accounted for by pure sinusoids (i.e., by a purely deterministic signal). Yet, often the "white noise" hypothesis is suspect. Furthermore, in sensor arrays the hypothesis of mutual couplings and local effect of scatterers suggests the presence of noise with ***short range***

correlation structure (e.g., the analog of say, $MA(1)$ or $MA(2)$ in time-series). In an effort to address such practical issues, in [22, 11] we develop canonical decompositions of second order statistics accounting for noise with short-range correlations. Such problems are naturally formulated as semi-definite programs and efficiently solved with existing software.

We wish to acknowledge the influence of early critique by R.E. Kalman, in the context of econometric "error-in-variables" models, on this line of research. Indeed, ordinary least-squares often lead to ill-posed solutions (as most clearly demonstrated in the econometric Frisch-Reiersøl problem [29]). The motivation behind [22] has been to address the case where information is available regarding noise statistics (typified by mutual couplings and interference in sensor arrays) and employ the system theoretic maxim that a maximal set of dependences is to be sought. The essence of our research has been to decompose second order statistics into a sum which reflects noisy and deterministic components. Accordingly, the decomposition is canonical and/or minimal in a suitable sense (see [22]). The formalism in [22] applies to the setting of distributed sensor arrays.

3.2 Distances between power spectra

Despite the centrality of spectral analysis in a wide range of scientific disciplines, no agreement exists as to what an appropriate distance measure between spectral density functions is. Some of the key contenders have been the Bregman distances, the Kullback-Leibler-von Neumann distance, the Itakura-Saito distance, and finally Battacharrya and Mahalanobis-type variants. Certain of these distances have a definite relevance when used to discriminate between two probability density functions. Yet ***none*** has any meaningful interpretation when applied to power spectra.

We present a new distance measure between power spectral densities and in fact, a (pseudo-) metric, which has a clear interpretation rooted in prediction theory. This is based on [24, 25].

Our starting point is to consider the degradation of the variance of the prediction error when the design of the predictor is based on the wrong choice among two alternatives. More specifically, let f_1, f_2 represent spectral density functions of discrete-time zero-mean stochastic processes $u_{f_i}(k)$ ($i \in \{1,2\}$ and $k \in \mathbb{Z}$), and let $p_{f_i}(\ell)$ ($\ell \in \{1,2,3,\ldots\}$) represent values for the coefficients that minimize the linear prediction error variance

$$\mathcal{E}\{|u_{f_i}(0) - \sum_{\ell=1}^{\infty} p(\ell) u_{f_i}(-\ell)|^2\}.$$

Thus, the optimal set of coefficients depends on the power spectral density function of the process, a fact which is acknowledged by the subscript in the notation $p_{f_i}(\ell)$. It is reasonable to consider as a distance between f_1 and f_2 the degradation of predictive error variance when the coefficients $p(\ell)$ are selected assuming one of the two, and then used to predict a stochastic process corresponding to the other spectral density function. The ratio of the "degraded" predictive error variance over the optimal error variance

$$\rho(f_1, f_2) := \frac{\mathcal{E}\{|u_{f_1}(0) - \sum_{\ell=1}^{\infty} p_{f_2}(\ell) u_{f_1}(-\ell)|^2\}}{\mathcal{E}\{|u_{f_1}(0) - \sum_{\ell=1}^{\infty} p_{f_1}(\ell) u_{f_1}(-\ell)|^2\}}$$

turns out to be equal to the ratio of the arithmetic over the geometric means of the fraction of the two spectral density functions, see [24, 25]. Then, the logarithm $\log \rho(f_1, f_2) =: \delta(f_1, f_2)$ represents a measure of dissimilarity between the "shapes" of f_1 and f_2 and, can be viewed, as analogous to "divergences" of Information Theory (such as the Kullback-Leibler relative entropy). Indeed,

$$\delta(f_1, f_2) = \log\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f_1(\theta)}{f_2(\theta)} \frac{d\theta}{2\pi}\right) - \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left(\frac{f_1(\theta)}{f_2(\theta)}\right) \frac{d\theta}{2\pi} \tag{3.8}$$

vanishes only when f_1/f_2 is constant on $[-\pi, \pi]$ and is positive otherwise. Considering the distance $\delta(f, f+\Delta)$ between a nominal power spectral density f and a perturbations $f + \Delta$, eliminating cubic terms and beyond, leads (modulo a scaling factor of 2) to the Riemannian pseudo-metric

$$g_f(\Delta) := \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\frac{\Delta(\theta)}{f(\theta)}\right)^2 \frac{d\theta}{2\pi} - \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\Delta(\theta)}{f(\theta)} \frac{d\theta}{2\pi}\right)^2 \tag{3.9}$$

on density functions. It was a pleasant surprise that, geodesic paths f_τ ($\tau \in [0,1]$) connecting spectral densities f_0, f_1 and having minimal length $\int_0^1 \sqrt{\delta(f_\tau, f_{\tau+d\tau})} = \int_0^1 \sqrt{g_{f_\tau}(\frac{\partial f_\tau}{\partial \tau})} d\tau$, can be explicitly computed

[24]. These turn out to be logarithmic intervals (also referred to as exponential families),

$$f_\tau(\theta) = f_0^{1-\tau}(\theta) f_1^\tau(\theta) \text{ for } \tau \in [0,1], \tag{3.10}$$

between the two extreme points. Furthermore, the length along such geodesics can be *explicitly* computed in terms of end points

$$d_g(f_1, f_2) := \sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\log\frac{f_1(\theta)}{f_2(\theta)}\right)^2 \frac{d\theta}{2\pi} - \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\frac{f_1(\theta)}{f_2(\theta)}\frac{d\theta}{2\pi}\right)^2}. \tag{3.11}$$

I.e., it is the "standard-deviation" of the difference $\log(f_1) - \log(f_2)$. This is a *pseudo-metric*. The "pseudo" refers only to the fact that it does not account for *constant* multiplicative factors.

It is rather interesting to point out that the $f \mapsto \log(f)$ maps power spectral densities onto a Euclidean space where quadratic norms such as (3.11) have a clear interpretation. In fact, with respect to the Riemannian metric (3.9) that we introduced, the space has zero curvature since geodesics are "logarithmic" straight lines. From this vantage point one may also consider alternative norms such as $\|\log(\frac{f_1}{f_2})\|_2$, etc. though without yet a natural interpretation.

It is interesting to compare the differential structure on power spectral density functions that we introduced above with the corresponding differential structure of "Information Geometry." In Information Geometry $f(\theta)$ corresponds to a probability density on $[-\pi, \pi]$ and the natural Riemannian metric is the *Fisher information metric* is (cf. [1, page 28]) which can be expressed in our framework as

$$g_{\text{Fisher},f}(\Delta) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\Delta(\theta)^2}{f(\theta)}\frac{d\theta}{2\pi} \tag{3.12}$$

(with $\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\frac{d\theta}{2\pi} = 1$ and $\frac{1}{2\pi}\int_{-\pi}^{\pi}\Delta(\theta)\frac{d\theta}{2\pi} = 0$ since both f, $f+\Delta$ need to be probability densities). Direct comparison reveals that the powers of $f(\theta)$ in (3.9) and (3.12) are different. Thus, it is curious and worth underscoring that in either differential structure, geodesics and geodesic lengths can be computed. For completeness we note that Information Geometry is a vast subject, originating in the work of Rao, Amari, Cencov and others, with a large following directed towards analogous geometric interpretations in Quantum theory. The starting point of

Information Geometry may be considered, in a way analogous to our development, to be the degradation of coding efficiency when the wrong choice between two probability distributions f_1 and f_2 is assumed. This degradation is precisely the Kullback-Leibler distance between the two, which can then give rise to the Riemannian metric $g_{\mathrm{Fisher},f}(\Delta)$, in a way analogous to our construction of $g_f(\Delta)$ from $\delta(\cdot,\cdot)$.

The idea to employ the degradation of performance with regard to specific tasks when the wrong choice between alternatives is used extends easily to a variety of contexts, and should be useful for that purpose as well. An alternative paradigm can be built on smoothing problems which we take up next but with a different goal in mind, namely, to provide an alternative to the maximum entropy principle.

The maximum entropy principle, as it is often invoked in time-series analysis ([28]), suggests the selection of a power spectrum which is consistent with autocorrelation data and corresponds to a random process least predictable from past observations. While this is a reasonable dictum when one is interested in prediction, it is often used regardless of the specific intent for the sought spectrum. The point we wish to raise becomes apparent when considering the relevance of another dictum, equally pertinent, albeit based on smoothing instead of prediction.

The variance $\mathcal{E}\{|u(0)-\sum_{\ell=1}^{\infty} p_f(\ell)u(-\ell)|^2\}$ of the optimal one-step-ahead (linear) predictor $\hat{u}(0|\text{past}) := \sum_{\ell=1}^{\infty} p_f(\ell)u(-\ell)$ is the *geometric mean* (see [26, page 183]) of f, i.e.,

$$\mathcal{E}\{|u(0)-\hat{u}(0|\text{past})|^2\} = m_{0,f} := \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left(f(\theta)\right) d\theta\right).$$

This is the content of the celebrated Kolmogorov-Szegö theorem. The entropy rate [27] is then defined as the negative integral of the logarithm of f (i.e., as $-\int \log(f(\theta))d\theta$). The notation $m_{0,f}$, taken from [2, page 23] for the *geometric mean*, is sought to contrast with the expression for the variance of the error

$$\mathcal{E}\{|u(0)-\hat{u}(0|\text{past}+\text{future})|^2\} = m_{-1,f} := \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)^{-1} d\theta\right)^{-1}$$

for the optimal smoothing filter $\hat{u}(0|\text{past}+\text{future}) := \sum_{\ell\neq 0}^{\infty} q_f(\ell)u(-\ell)$. This expression has been derived in [23] where it was also noted that it represents the *harmonic mean* of f. (Naturally, the harmonic mean is always $\leq$ to the geometric mean, since smoothing uses more data.)

Applications abound where records need to be interpolated, or where the indexing of data collected via a sensor array represents spatial ordering and not time ordering. In all such applications there is no natural "time-arrow" and, hence, it is imperative that Burg's maximum entropy principle is re-evaluated.

Thus, in the context of time series analysis, both Burg's "predictive entropy" $-\int \log(f(\theta))d\theta$ and the "smoothing entropy" $\int f(\theta)^{-1}d\theta$ derived in [23] in work supported by the present grant, relate to the level of unpredictability in these two vastly different situations. Burg's entropy has been also used as a regularizing functional in inverse problems (see [9]). But the latter functional can be used equally well for similar modeling purposes. For instance, we have shown in [23] that extremal spectra with respect to the second choice give rise to all-pole Markovian models very much like Burg's maximum entropy AR-models, but with one important difference. The poles in these models appear with fractional powers. Such fractional powers are often encountered in processes with long "memory."

There is an apparent dichotomy between what a deterministic process is, depending on whether we consider a one-sided or a two-sided past. Stationary time-series are said to be deterministic in the Kolmogorov sense if $\log(f) \notin L_1$. When we consider determinism with respect to a two-sided past, then the corresponding condition weakens to $f^{-1} \notin L_1$, because it is only then that the smoothing error is zero. This dichotomy raises similar questions for spatial processes and fields. This is especially pertinent for applications where space-time data are collected via sensor arrays.

Ever since the early days of statistical mechanics, entropy has been a very elusive concept. Yet, from a mathematical and computational standpoint, entropy and entropy-rate functionals can be thought of as natural barriers of convex sets and positive cones (i.e., of probability simplices, or of cones of spectral density functions). They thus can be used to identify solutions to ill-posed inverse problems. The history of such a viewpoint and of earlier developments can be looked at in the PI's recent publication [9]. Besides the usual entropy functionals introduced by Shannon-von Neuman, Burg, and Kullback-Leibler-Linblad-Leib, discussed in [9] there is a plethora of alternatives, such as the Rényi entropy, Tsallis entropy, and more recently generalized means in the PI's work [24]. This last work provides, as discussed earlier, notion of distance which are compatible with prediction problems and are suitable for incorporating uncertainty into correlation measurements (in the spirit of [8]).

3.3 Software

The following is only a brief overview of representative routines used for high resolution spectral analysis of *scalar* time-series. Additional information and Matlab code is provided at
http://www.ece.umn.edu/users/georgiou/files/reports.html.
Please contact the principal investigator at tryphonumn.edu for comments/input and for subsequent releases of Matlab-based code for high resolution spectral analysis. We briefly explain how the software we developed can be used to resolve sinusoids —a benchmark problem.

We begin we time-domain data and some prior information as to the frequency range of interest. A filter-bank (one input many outputs) is then selected with a bandpass characteristic over the frequency range of interest. It consists of a dynamical system

$$x_{k+1} = Ax_k + Bu_k$$

with A, B matrices of size $n \times n$, and $n \times m$ respectively. Of course, when the time-series u_k is scalar, $m = 1$. The time-series is considered at present to have zero mean and is adjusted accordingly. An observation record

$$\{u_1, u_2, \ldots, u_N\}$$

is typically available, and on the basis of that an estimate of the state-covariance

$$P = E\{x_k x_k^*\}$$

is obtained using routine `dlsim_complex.m`. Relevant theory and a number of routines, e.g., the ones below,

Name	usage
`sm.m`	`[fr_lines,ampl_lines]=sm(P,A,B,k)`
`me.m`	`me_spect=me(P,A,B,omega)`
`envlp.m`	`env=envlp(P,A,B,omega,noiselevel)`

http://www.ece.umn.edu/users/georgiou/files/reports.html.

can be used for spectral analysis. The examples above determine (a) spectral lines consistent with the data (`sm.m`), (b) a candidate spectrum for u_k which is consistent with the data P and is of maximal entropy (`me.m`), and (c) an envelop for the amplitude of all consistent with the data spectral lines (`envlp.m`).

The resolution of the above routines strongly depends on the choice of A, B and on the variance of the estimator for the state covariance P. Tradeoffs between robustness and resolution using such methods is the subject of a Ph.D. thesis by A.Nasiri-Amini (December 2005) and discussed in [32]. These provide guidelines for optimal design of *input-to-state* filters and theoretical bounds for the expected gains in resolution.

In practice, as a rule of thumb, there are two parameters that dictate the performance of the relevant spectral estimators: the time-constant of the *input-to-state* filter

$$G(z) = (zI - A)^{-1}B$$

and its bandpass character. Routines `cjordan2.m` and `mjordan.m` can be used for designing suitable (A, B) pairs for scalar and vectorial time-series, respectively. Typically one needs only specify a (complex) eigenvalue(s) for A and the size of the corresponding Jordan block(s). The modulus of the eigenvalues dictates the time-constant of $G(z)$ and the phase specifies the band pass character. Finally, the pair is then normalized to satisfy

$$AA^* + BB^* = I$$

where I is the identity, for numerical reasons. Typically, A can be chosen to have one Jordan block (when u_k is scalar) or as a Kronecker product of such a matrix with the identity (as in `mjordan.m`).

Routine `demo1.m` exemplifies the performance of the above for an academic example of separating two sinusoids in background noise. Because of the band-pass character of $G(s)$ and the fact that the framework relies on the state-covariance of $G(s)$, the performance of all the above is impervious to color noise (as long as it is relatively white over the passband of $G(s)$). Yet, in the example we use white noise for simplicity. Figure 3.2 displays a typical output. The "true" spectrum of the time-series u_k is represented by the red "noise level" and two red arrows for the sinusoids at frequencies 1.3 and 1.35 [rad/sec]. The subplots on the right represent "zoom-in" of those on the left, focusing on a frequency range of interest $[1.2, 1.5]$. The data record contains only 100 points. The fourier-based reconstruction is shown in green in subplots (2,1) and, magnified in (2,2). It is apparent that the length of the data record, which is very short compared to the separation of the two spectral lines, does not permit periodogram based techniques to achieve any level of resolution. Our high resolutions methods provide very accurate spectral estimates. These are shown in blue. In particular, estimated spectral lines are indicated with blue arrows and are compared with the actual spectral lines of the signal in subplots (1,1) and in (1,2). Subplots (3,1) and (3,2) show the computed envelop that bounds all power spectra which are consistent with the estimated generalized statistics. For comparison we mark the position of the actual frequencies of the two sinusoids with thin red lines. Finally, subplots (4,1) and (4,2) show the maximum entropy power spectrum which is consistent with the generalized statistics. These three alternatives (blue arrows in the first row, and blue envelop and spectrum in rows 3 and 4) have been produced with the above routines. They indicate a remarkable consistency and accuracy which has been confirmed by theoretical calculations and error bounds ([32], [31], [33]).

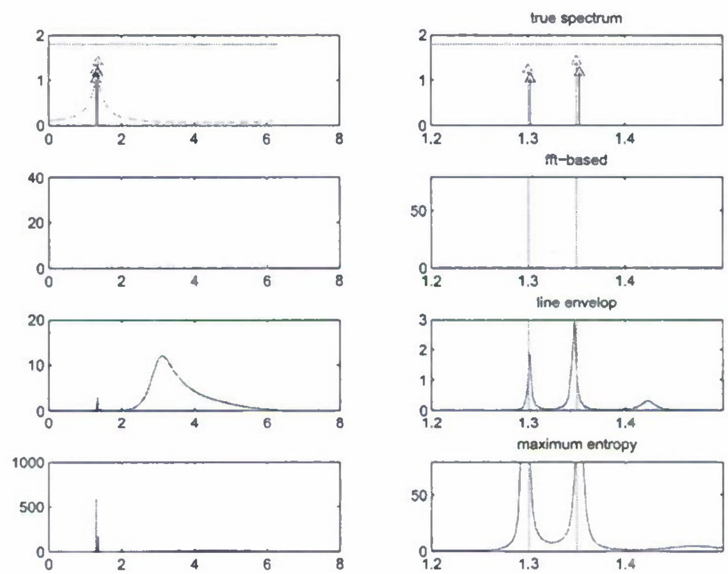

Figure 3.2: Original and reconstructed power spectra

Bibliography

[1] S. Amari, **Differential-geometrical methods in statistics**, Lecture notes in statistics, Springer-Verlag, Berlin, 1985.

[2] E.F. Beckenbach and R. Bellman, **Inequalities**, Springer-Verlag, Berlin-Heidelberg, 198 pages,1965.

[3] C. I. Byrnes, T.T. Georgiou, and A. Lindquist, "A generalized entropy criterion for rational Nevanlinna-Pick interpolation: A convex optimization approach to certain problems in systems and control," *IEEE Trans. on Automatic Control*, **45(6):** 822-839, June 2001.

[4] C. I. Byrnes, T.T. Georgiou, and A. Lindquist, "A new approach to spectral estimation: A tunable high-resolution spectral estimator," *IEEE Trans. on Signal Proc.*, **48(11):** 3189-3206, November 2000.

[5] C.I. Byrnes, T.T. Georgiou, A. Lindquist, and A. Megretski, "Generalized interpolation in H^∞ with a complexity constraint," *Trans. of the American Math. Society*, (electronically published on December 9, 2004) *Trans. Amer. Math. Soc.* **358(3)**: 965-987, 2006.

[6] T.T. Georgiou and A. Lindquist, "Kullback-Leibler approximation of spectral density functions," *IEEE Trans. on Information Theory*, **49(11)**, November 2003.

[7] T.T. Georgiou, "Signal Estimation via Selective Harmonic Amplification: MUSIC redux," *IEEE Trans. on Signal Processing*, **48(3)**: 780-790, March 2000.

[8] T.T. Georgiou, "Solution of the general moment problem via a one-parameter imbedding," *IEEE Trans. on Automatic Control*, **50(6)**: 811-826, June 2005.

[9] T.T. Georgiou, "Relative Entropy and the multi-variable multi-dimensional Moment Problem," *IEEE Trans. on Information Theory*, **52(3):** 1052 - 1066, March 2006.

[10] T.T. Georgiou and A. Lindquist, "Remarks on control design with degree constraint," *IEEE Trans. on Automatic Control*, **51(7):** 1150-1156, July 2006.

[11] T.T. Georgiou, "Decomposition of Toeplitz matrices via convex optimization," *IEEE Signal Processing Letters*, **13(9):** 537- 540, September 2006.

[12] T.T. Georgiou, "A differential equation for the LMI feasibility problem," submitted to *IEEE Trans. on Automatic Control, under revision.*

[13] T.T. Georgiou, " Toeplitz covariance matrices and the von Neumann relative entropy," in Control and Modeling of Complex Systems: Cybernetics in the 21st Century: Festschrift volume for Professor H. Kimura; K. Hashimoto, Y. Oishi, and Y. Yamamoto, Eds. Boston, MA: Birkhauser, 2003.

[14] T.T. Georgiou, " Structured covariances and related approximation questions," in Lecture Notes in Control and Information Sciences, Volume 286, 2003, pp. 135-140, Springer Verlag, 2003.

[15] T.T. Georgiou, " The mixing of state covariances," in Lecture Notes in Control and Information Sciences, Volume 289, 2003, pp. 207-212, Springer Verlag, 2003.

[16] T.T. Georgiou and A. Lindquist, "Two Alternative Views on Control Design with Degree Constraint," Proc. 44th IEEE Conference on Decision and Control, December 2005, pp. 3645-3650.

[17] T.T. Georgiou, "Relative Entropy and Moment Problems," Proc. 44th IEEE Conference on Decision and Control, December 2005, pp. 4397 - 4403.

[18] T.T. Georgiou, "Singular decomposition of state covariances," Proceedings of the 2006 American Control Conference, 6 pp., June 2006.

[19] T.T. Georgiou, " Spectral Estimation via Selective Harmonic Amplification, *IEEE Trans. on Automatic Contr.*," January 2001, **46(1)**: 29-42.

[20] T.T. Georgiou, "The structure of state covariances and its relation to the power spectrum of the input," *IEEE Trans. on Automatic Control*, **47(7)**: 1056-1066, July 2002.

[21] T.T. Georgiou, "Spectral analysis based on the state covariance: the maximum entropy spectrum and linear fractional parameterization," *IEEE Trans. on Automatic Control*, **47(11)**: 1811-1823, November 2002.

[22] T.T. Georgiou, "The Caratheéodory-Fejér-Pisarenko decomposition and its multivariable counterpart," preprint at: (http://arxiv.org/abs/math.PR/0509225); to appear in *IEEE Trans. on Automatic Control*, March 2007.

[23] T.T. Georgiou, "The maximum entropy ansatz in the absence of a time-arrow: fractional pole models," preprint at: (http://arxiv.org/abs/math/0601648/), submitted to the *IEEE Trans. on Information Theory*, January 2006.

[24] T.T. Georgiou, "Distances between power spectral densities," preprint at: (http://arxiv.org/abs/math/0607026), submitted to the *IEEE Trans. on Signal Processing*, July 2006.

[25] T.T. Georgiou, "An intrinsic metric for power spectral density functions," preprint at: (http://arxiv.org/abs/math/0608486), submitted to the *IEEE Signal Processing Letters*, August 2006.

[26] U. Grenander and G. Szegö, **Toeplitz Forms and their Applications**, Chelsea, 1958.

[27] S. Haykin, **Nonlinear Methods of Spectral Analysis**, Springer-Verlag, New York, 247 pages, 1979.

[28] E.T. Jaynes, "On the rationale of maximum entropy methods," *Proc. IEEE*, **70**: 939-952, 1982.

[29] R.E. Kalman, "Identification from real data," **Current Developments in the Interfact: Econmics, Econometrics, Mathematics**, Reidel, Dordrecht, pp. 161-196, 1982.

[30] J. Karlsson and T.T. Georgiou, "Signal analysis, moment problems & uncertainty measures," Proc. 44th IEEE Conference on Decision and Control, December 2005, pp. 5710 - 5715.

[31] A. Nasiri Amini, E. Ebbini, and T.T. Georgiou, "Noninvasive tissue temperature estimation using high-resolution spectral analysis techniques," *IEEE Trans. on Biomedical Engineering*, **52(2)**: 221-228, February 2005.

[32] A. Nasiri Amini and T.T. Georgiou, "Tunable Spectral Estimators Based on State-Covariance Subspace Analysis," *IEEE Trans. on Signal Processing,* **54(7):** 2662-2671, July 2006.

[33] A. Nasiri Amini, "Tradeoffs in Resolution of Nonlinear Spectral Estimation," PhD thesis, University of Minnesota, December 2005.

[34] A. Nasiri Amini, E. Ebbini, and T. Georgiou, "Noninvasive tissue temperature estimation via state-covariance spectral estimation," *Proc. 10th IEEE DSP workshop,* Taos, NM, August 2004.

[35] A. Nasiri Amini and T.T. Georgiou, "SAR imaging via state-covariance subspace methods," under revision.

[36] A. Nasiri Amini and T.T. Georgiou, "Avoiding Ambiguity in Beamspace Processing," *IEEE Signal Processing Letters,* **12(5)**: 372 - 375, May 2005.

[37] A. Nasiri Amini, S. Takyar, and T.T. Georgiou, "A Homotopy Approach for Multirate Spectrum Estimation," Proceedings of the 2006 IEEE International Conf. on Acoustics, Speech and Signal Processing, vol. 3, pages 532-535, May 2006.

[38] P. Stoica and R. Moses, **Introduction to Spectral Analysis**, Prentice Hall, 2005.

[39] S. Takyar, A. Nasiri Amini, and T.T. Georgiou, "Sensitivity shaping with degree constraint via semidefinite programming," Proceedings of the 2006 American Control Conference, 4 pp., June 2006.

[40] S. Takyar, A. Nasiri Amini, and T.T. Georgiou, "Spectral analysis from multi-rate observations: Itakura-Saito based approximation of power spectra and Capon-like spectral envelopes," Proceedings of the Fourth IEEE Workshop on Sensor Array and Multi-channel Processing (SAM-2006), Waltham, Massachusetts, July 2006

[41] United States. Patent 6,400,310, "A tunable high-resolution spectral estimator," (joint co-inventors with C.I. Byrnes and A. Lindquist), June 4, 2002.

[42] H.L. Van Trees, **Optimum Array Processing: part IV of Detection, Estimation and Modulation Theory**, Wiley-Interscience, 2002.

[43] S.R.S. Varadhan, **Probability Theory**, AMS, 2000.

[44] S. Varigonda, T.T. Georgiou, and P. Daoutidis, "Numerical solution of the optimal periodic control problem using differential flatness," *IEEE Trans. on Automatic Control*, **49(2)**:271 - 275, February 2004.

Chapter 4

Principal investigator vitae

Tryphon T. Georgiou

Education	*Doctor of Philosophy in Electrical Engineering,* **University of Florida**, 1983. *Diploma in Electrical and Mechanical Engineering,* **NTUA, Greece**, 1979.
Positions	*Vincentine Hermes-Luh Chair.* Dept. of ECE, *Univ. of Minnesota*, 2002–present. *Professor. University of Minnesota*, 1994–present. *Associate Prof. University of Minnesota*,1989–1994. *Associate Prof. Iowa State University*, 1988–1989. *Assistant Prof. Iowa State University*, 1986–1988. *Assistant Prof. Florida Atlantic University*, 1983–1986.
Posts	*Member of Board of Governors, IEEE Control Systems Society*, elected, 2002–2004. *Associate Editor*, SIAM Journal on Control and Optimization, 1988–1995. *Associate Editor*, Systems and Control Letters, 1995–present. *Associate Editor*, IEEE Transactions on Automatic Control, 1991–1992. *Co-Director.* Control Science and Dyn. Systems Center. *Univ. of Minn.*, 1990–present.

4.1 Awards

1. **Fellow of the IEEE**, January 2000, for contributions to the theory of robust control.

2. **2003 IEEE Control Systems Society G.S. Axelby Outstanding Paper Award**[4],
   for the paper: C. Byrnes, T.T. Georgiou, and A. Lindquist, "A generalized entropy criterion for Nevanlinna-Pick interpolation: A convex optimization approach to certain problems in systems and control", *IEEE Trans. on Automatic Control*, **45(6)**: 822-839, June 2001.

3. **1999 IEEE Control Systems Society G.S. Axelby Outstanding Paper Award**[1],
   for the paper: T.T. Georgiou and M.C. Smith "Robustness Analysis of Nonlinear Feedback Systems: An Input-Output Approach," *IEEE Trans. on Automatic Contr.*, 42(9): 1200-21, 1997.

4. **1992 IEEE Control Systems Society G.S. Axelby Outstanding Paper Award**[4],
   for the paper: T.T. Georgiou and M.C. Smith "Optimal robustness in the gap metric",
   IEEE Trans. on Automat. Contr., 35(6): 673-86, 1990.

5. **1993-1994 Best Instructor Award**. Presented by the Institute of Technology Student Board,
   University of Minnesota, May 1994.

4.2 Patents

1. Method and Apparatus for a tunable high-resolution spectral estimator. Co-discoverers: Chris Byrnes (Washington University) and Anders Lindquist (Royal Institute of Technology, Sweden). ***U.S. Patent No. 6,400,310***. International patent pending.

[1]The G.S. Axelby award is for *"outstanding papers published in the IEEE Transactions on Automatic Control during the previous two calendar years before the year of the award, based on originality, clarity, potential impact on the theoretical foundations of control, and practical significance in application"*.

4.3 Personnel partially supported by the grant

1. Dr. Tryphon T. Georgiou, Professor, University of Minnesota
2. Dr. Ali Nasiri Amini, former research assistant and doctoral student, currently wiith America Online Inc.
3. Mr. Johan Karlsson, visiting scholar, currently with Royal Institute of Technology, Sweden
4. Mr. Shahrouz Takyar, doctoral student, Univ. of Minnesota

4.4 Recent Service & Synergistic Activities

1. Member of Steering & Program committees for the International Symposium on Mathematical Theory of Networks and Systems (MTNS), and member of Organizing committee for MTNS 2008.
2. Member of Technical program committees for the IEEE International Conference on Decision and Control 2006 and the European Control Conference 2007.
3. Member of the University of Minnesota Faculty Senate and Co-director of the Control and Dynamical Systems Center at the University of Minnesota (2005-2007).

4.5 AFRL point of contact

Dan Herrick C Civ AFRL/DESA Kirtland AFB
Tel: 505-853-5189, Fax: 505-853-1698 FAX, Secretary: 505-846-4016

4.6 Publications on research under AFOSR AF/F49620-03-1-0120

4.6.1 Journal publications

1. T.T. Georgiou, "Spectral analysis based on the state covariance: the maximum entropy spectrum and linear fractional parameterization," *IEEE Trans. on Automatic Control*, **47(11):** 1811-1823, November 2002.

2. T.T. Georgiou and A. Lindquist, "Kullback-Leibler approximation of spectral density functions," *IEEE Trans. on Information Theory*, **49(11)**, November 2003.

3. S. Varigonda, T.T. Georgiou, and P. Daoutidis, "Numerical solution of the optimal periodic control problem using differential flatness," *IEEE Trans. on Automatic Control*, **49(2)**:271 - 275, February 2004.

4. A. Nasiri Amini, E. Ebbini, and T.T. Georgiou, "Noninvasive tissue temperature estimation using high-resolution spectral analysis techniques," *IEEE Trans. on Biomedical Engineering*, **52(2)**: 221-228, February 2005.

5. T.T. Georgiou, "Solution of the general moment problem via a one-parameter imbedding," *IEEE Trans. on Automatic Control*, **50(6)**: 811-826, June 2005.

6. A. Nasiri Amini and T.T. Georgiou, "Avoiding Ambiguity in Beamspace Processing," *IEEE Signal Processing Letters*, **12(5)**: 372 - 375, May 2005.

7. C.I. Byrnes, T.T. Georgiou, A. Lindquist, and A. Megretski, "Generalized interpolation in H^∞ with a complexity constraint," *Trans. of the American Math. Society*, (electronically published on December 9, 2004) *Trans. Amer. Math. Soc.* **358(3)**: 965-987, 2006.

8. T.T. Georgiou, "Relative Entropy and the multi-variable multi-dimensional Moment Problem," *IEEE Trans. on Information Theory*, **52(3)**: 1052 - 1066, March 2006.

9. T.T. Georgiou and A. Lindquist, "Remarks on control design with degree constraint," *IEEE Trans. on Automatic Control*, **51(7)**: 1150-1156, July 2006.

10. A. Nasiri Amini and T.T. Georgiou, "Tunable Spectral Estimators Based on State-Covariance Subspace Analysis," *IEEE Trans. on Signal Processing, to appear.*

11. T.T. Georgiou, "Decomposition of Toeplitz matrices via convex optimization," *IEEE Signal Processing Letters*, **13(9)**: 537- 540, September 2006.

12. T.T. Georgiou, "The Carathéodory-Fejér-Pisarenko decomposition and its multivariable counterpart," *IEEE Trans. on Automatic Control,* to appear, March 2007 preprint http://arxiv.org/abs/math/0509225.

13. T.T. Georgiou, "The maximum entropy ansatz in the absence of a time-arrow: fractional pole models," *IEEE Trans. on Information Theory*, under review, January 2006, preprint: http://arxiv.org/abs/math/0601648.

14. T.T. Georgiou, "Distances between power spectral densities," *IEEE Trans. on Signal Processing, under review*, July 2006, preprint: http://arxiv.org/abs/math/0607026.

15. T.T. Georgiou, "An intrinsic metric for power spectral density functions," *IEEE Signal Processing Letters,* to appear, 2007, preprint: http://arxiv.org/abs/math/0608486.

4.6.2 Book chapters

16. T.T. Georgiou, P.J. Olver, and A. Tannenbaum, Maximal entropy for reconstruction of back projection images, "IMA Volumes in Mathematics and its Applications," *Volume 133: Mathematical methods in computer vision* Springer-Verlag, New York, 2002.

17. T.T. Georgiou, Toeplitz covariance matrices and the von Neumann relative entropy, in Control and Modeling of Complex Systems: Cybernetics in the 21st Century: Festschrift volume for Professor H. Kimura; K. Hashimoto, Y. Oishi, and Y. Yamamoto, Eds. Boston, MA: Birkhauser, 2003.

18. T.T. Georgiou, Structured covariances and related approximation questions, in Lecture Notes in Control and Information Sciences, Volume 286, 2003, pp. 135-140, Springer Verlag, 2003.

19. T.T. Georgiou, The mixing of state covariances, in Lecture Notes in Control and Information Sciences, Volume 289, 2003, pp. 207-212, Springer Verlag, 2003.

4.6.3 Theses

20. A. Nasiri Amini, "Tradeoffs in Resolution of Nonlinear Spectral Estimation," PhD thesis, University of Minnesota, December 2005.

4.6.4 Refereed Conference Publications

21. A. Nasiri Amini and T.T. Georgiou, "Statistical analysis of state-covariance subspace-estimation methods," Proceedings of the 41st IEEE Conference on Decision and Control, pp. 2633 -2638, December 2002.

22. T.T. Georgiou and A. Lindquist, "Kullback-Leibler approximation of spectral density functions," Proceedings of the 42st IEEE Conference on Decision and Control, December 2003.

23. A. Nasiri Amini, E. Ebbini, and T. Georgiou, "Noninvasive tissue temperature estimation via state-covariance spectral estimation," *Proc. 10th IEEE DSP workshop,* Taos, NM, August 2004.

24. T.T. Georgiou and A. Lindquist, "Two Alternative Views on Control Design with Degree Constraint," Proc. 44th IEEE Conference on Decision and Control, December 2005, pp. 3645-3650.

25. Johan Karlsson and Tryphon T. Georgiou, "Signal analysis, moment problems & uncertainty measures," Proc. 44th IEEE Conference on Decision and Control, December 2005, pp. 5710 - 5715.

26. Tryphon T. Georgiou, "Relative Entropy and Moment Problems," Proc. 44th IEEE Conference on Decision and Control, December 2005, pp. 4397 - 4403.

27. Tryphon T. Georgiou, "Singular decomposition of state covariances," Proceedings of the 2006 American Control Conference, 6 pp., June 2006.

28. Mir Shahrouz Takyar, Ali Nasiri Amini, and Tryphon T. Georgiou, "Sensitivity shaping with degree constraint via semidefinite programming," Proceedings of the 2006 American Control Conference, June 2006.

29. Mir Shahrouz Takyar, Ali Nasiri Amini, and Tryphon T. Georgiou, "Spectral analysis from multi-rate observations: Itakura-Saito based approximation of power spectra and Capon-like spectral envelopes," Proceedings of the Fourth IEEE Workshop on Sensor Array and Multi-channel Processing (SAM-2006), Waltham, Massachusetts, July 2006

30. Ali Nasiri Amini, Shahrouz Takyar, and Tryphon Georgiou, "A Homotopy Approach for Multirate Spectrum Estimation," Proceedings of the 2006 IEEE International Conf. on Acoustics, Speech and Signal Processing.

31. Tryphon T. Georgiou, "The error variance of the optimal linear smoother is the harmonic mean of the power spectral density," Proc. 45th IEEE Conference on Decision and Control, December 2006.

32. Johan Karlsson, Tryphon T. Georgiou, and A. Lindquist, "The Inverse Problem of Analytic Interpolation with Degree Constraint," Proc. 45th IEEE Conference on Decision and Control, December 2006.

33. Shahrouz Takyar, Ali Nasiri Amini, and Tryphon Georgiou, "Weight selection in Interpolation with a Dimensionality Constraint," Proc. 45th IEEE Conference on Decision and Control, December 2006.

4.7 Acknowledgment/Disclaimer

This work was sponsored (in part) by the Air Force Office of Scientific Research, USAF, under grant/contract number AF/F49620-03-1-0120. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.